

Register No.: ..... Name: .....

## SAINTGITS COLLEGE OF ENGINEERING (AUTONOMOUS)

(AFFILIATED TO APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY, THIRUVANANTHAPURAM)

FOURTH SEMESTER MBA DEGREE EXAMINATION (Regular), JULY 2022

(2020 Scheme)

Course Code : 20MBA442

Course Name: Financial Applications for Machine Learning

Max. Marks : 60

Duration: 3 Hours

### PART A

(Answer all questions. Each question carries 2 marks)

1. Distinguish between overfitting and underfitting. How can it affect model generalization?
2. Describe the uses of decision tree for classification purpose
3. Define hyperplane in a Support Vector Machine
4. Differentiate between Training data and Testing Data
5. Define Market Basket Analysis

### PART B

(Answer any 3 questions. Each question carries 10 marks)

6. Explain the Linear discriminant Analysis and its elaborate a suitable application in finance Domain
7. For a Skin Cancer Dataset given below, construct a decision tree

Name	Skin Color	Height	Weight	Cosmetic Use	Class
Laxmi	Brown	Average	Light	No	Yes
Jimmy	Brown	Tall	Average	Yes	No
Samuel	White	Short	Average	Yes	No
Anita	Brown	Short	Average	No	Yes
Riaz	Black	Average	Heavy	No	Yes
Girish	White	Tall	Heavy	No	No
Sharukh	White	Average	Heavy	No	No
Shalini	Brown	Short	Light	Yes	No

8. Describe the significance of Kernel functions in SVM. [4 Marks]  
Describe any three Kernel functions. [6 Marks]
9. A. Describe the characteristics of the Back propagation Algorithm [5 Marks]  
B. Calculate the output  $y$  of a three-input neuron with bias. The input feature vector is  $(x_1, x_2, x_3) = (0.8, 0.6, 0.4)$  and weight values are  $[w_1, w_2, w_3, b] = [0.2, 0.1, -0.3, 0.35]$ . Use binary Sigmoid function as activation function. [5 Marks]
10. Describe How Principal Component Analysis can be used to reduce the dimensionality of the Data

**PART C**

*(Compulsory question, the question carries 20 marks)*

11. A) Cluster the following eight points A1(2,10), A2(2,5), A3(8,4), A4(5,8), A5(7,5), A6(8,4), A7(1,2), A8 (4,9) into three clusters using K-means algorithm. Assume the initial seed points of the Cluster to be A1, A4 and A7 respectively. Trace the K-means algorithm for 2 iterations. At the end of each iteration, show the new clusters and the centers of the new clusters. (15 Marks)
- B) Explain any two approaches to find optimal number of clusters (5 Marks)

\*\*\*\*\*